Molecular Psychiatry (2007) 12, 870–880

www.nature.com/mp

ORIGINAL ARTICLE

Evidence for the association of the *S100β* gene with low cognitive performance and dementia in the elderly

J-C Lambert[1], S Ferreira[1,2], J Gussekloo[3], L Christiansen[4], G Brysbaert[1], E Slagboom[3], D Cottel[1], T Petit[5], J-J Hauw[6], ST DeKosky[7], F Richard[1], C Berr[8], C Lendon[9], M Ilyas Kamboh[7], D Mann[10], K Christensen[4], R Westendorp[3] and P Amouyel[1]

[1]*INSERM U744, Institut Pasteur de Lille, Université de Lille 2, Lille, France;* [2]*Genoscreen, Lille, France;* [3]*Section of Gerontology and Geriatrics, Leiden University Medical Center, Leiden, The Netherlands;* [4]*Department of Epidemiology, Institute of Public Health, Epidemiology, University of Southern Denmark, Odense, Denmark;* [5]*Zoo de La Palmyre, Les Mathes, France;* [6]*INSERM IFR 70, APHP, GH Pïtié-Salpêtrière, Laboratoire de Neuropathologie R Escourolle, Paris, France;* [7]*Department of Human Genetics and Alzheimer's Disease Research Centre, University of Pittsburgh, Pittsburgh, PA, USA;* [8]*INSERM E 0361, University of Montpellier, Montpellier, France;* [9]*Department of Psychiatry, Neuroscience, University of Birmingham, Birmingham, UK and* [10]*Division of Medicine and Neurosciences, Greater Manchester Neurosciences Centre, University of Manchester, Manchester, UK*

Variations in the *S100β* gene may be instrumental in producing a continuum from mild cognitive decline to overt dementia. After screening 25 single nucleotide polymorphisms (SNPs) in *S100β*, we observed association of the rs2300403 intron 2 SNP with poorer cognitive function in three independent populations. Moreover, we detected a significant association of this SNP with increased risk of developing dementia or Alzheimer's disease (AD) in six independent populations, especially in women and in the oldest. Furthermore, we characterised a new primate-specific exon within intron 2 (the corresponding mRNA isoform was called S100β2). S100β2 expression was increased in AD brain compared with controls, and the rs2300403 SNP was associated with elevated levels of S100β2 mRNA in AD brains, especially in women. Therefore, this genetic variant in *S100β* increases the risk of low cognitive performance and dementia, possibly by favouring a splicing event increasing S100β2 isoform expression in the brain.
Molecular Psychiatry (2007) **12,** 870–880; doi:10.1038/sj.mp.4001974; published online 6 March 2007

Keywords: Alzheimer; dementia; S100β; expression; polymorphism; splicing

Introduction

In many individuals, an initial cognitive impairment inevitably and progressively leads to dementia, of which Alzheimer's disease (AD) is the principal cause. A continuum may exist ranging from slight and subtle memory loss through overt dementia.[1] If this is so, it is probable that shared environmental and genetic susceptibility factors may relate these two entities. With regard to dementia, three risk factors are well established: age, gender and family history, probably to be underpinned by genetic susceptibility factors.

Until now, family history has been decoded mainly through the study of early-onset inherited forms of AD accounting for less than 2% of the cases. Three genes have already been identified: amyloid precursor protein, presenilin 1 and presenilin 2.[2] However, none of these have been associated with cognitive decline in general ageing populations. The genetic determinants of the most common late-onset forms of the affection defined as sporadic (without obvious mendelian pattern of inheritance) appears to be far more complex. Numerous genetic susceptibility factors have been suspected. To date, only the apolipoprotein E gene polymorphism (*APOE*) has been consistently identified as a genetic determinant of AD.[3] In contrast to the early-onset AD genes, *APOE* has been associated with early cognitive impairment in ageing populations[4] and may constitute a common determinant of the suspected continuum.

Several lines of evidence prompted us to consider whether the S100 calcium-binding protein B (*S100β*) gene might also be implicated within this continuum. First, this gene is located on chromosome 21q22.3, in a chromosomal region consistently linked to familial late-onset forms of AD.[5] Second, S100βprotein exhibits cytotoxic and neurotrophic properties.[6] Third, S100β protein interacts with the microtubule-associated protein tau,[7] and is a component of amyloïd

Correspondence: Dr J-C Lambert, unité INSERM 744, Institut Pasteur de Lille, BP 245,1, rue du professeur Calmette, 59019 Lille cédex, France.
E-mail: jean-charles.lambert@pasteur-lille.fr
Received 5 October 2006; revised 7 November 2006; accepted 4 December 2006; published online 6 March 2007

plaques.[8] Finally, S100β protein has been claimed to be a peripheral marker of brain damage[9] and of cognitive decline following cardiac surgery.[10]

Thus, we undertook an in-depth genomic analysis of *S100β* locus to identify genetic mechanisms or variations of potential interest that may suggest this gene could play a role in the continuum from cognitive decline to dementia.

Materials and methods

Populations

The main characteristics of the populations are described in Tables S2 and S3 (Supplementary material). All subjects or, in those with substantial cognitive impairment, a caregiver, legal guardian or other proxy gave written informed consent for participation in this study. The study protocols for all populations were reviewed and approved by the appropriate institutional review boards of each country.

The ELDNOR study. The population study consisted of individuals aged between 60 and 105 years in a large representative sample ($n=1051$), living in retirement home, whose cognitive functions and health status were systematically registered in the French Elderly in the North (ELDNOR) study.[11] From this population-based study, a sub-population of 254 individuals was first used for the linkage disequilibrium (LD) estimation between the S100β single nucleotide polymorphisms (SNPs) and for the systematic screening of their impact on the cognitive performances (age $=80.6\pm8.5$ years, 33% of men, mini mental-state examination (MMSE) >24). The analysis of the SNPs showing a significant effect on cognitive performances in this first screening was next extended to 815 individuals exhibiting a MMSE >10. Furthermore, from this population-based study, two groups with highly contrasted levels of cognitive function were selected: 221 demented individuals, including all ELDNOR subjects with a MMSE score <10 or a diagnosis of dementia according to criteria given in DSM-III-R and a group of 146 control subjects, including all ELDNOR subjects with a MMSE >27 and no symptom of dementia. Eighteen demented cases were unsuccessfully genotyped for the rs2300403 SNP.

The Danish twin study. This study population is based on participants of the population-based nationwide survey 'The Longitudinal Study of Ageing Danish Twins' (LSADT), which includes Danish twins aged 70 years and older ($n=513$, with a MMSE >10).[12] The survey is an ongoing longitudinal study initiated in 1995 that comprises face-to-face interviews focusing on health and lifestyle issues, including assessment of MMSE. The subjects were randomly selected among those participants from the LSADT 1997 and 1999 waves for which both DNA and MMSE assessment were available. Only one twin from each pair was included in this analysis. In the Danish twin study, the demented status was not available.

The Leiden 85-plus study. The Leiden 85-plus study is a prospective population-based study that recruited elderly inhabitants of Leiden (The Netherlands).[13] People aged 85 years and above were enrolled between 1987 and 1989 or between 1997 and 1999. Cognitive functions were measured with the MMSE. Those with a MMSE-score below 24 points were psychiatrically examined to diagnose dementia following DSM-III criteria or diagnoses of dementia were gathered by reviewing medical records. DNA was available for 1218 participants (430 men and 938 women). The impact of the rs2300403 on cognitive performance was first assessed on 1049 individuals exhibiting a MMSE >10. Furthermore, from this population-based study, two groups with highly contrasting levels of cognitive functions were selected: 206 individuals with a diagnosis of dementia and 438 controls with a MMSE score above 27 and no symptoms of dementia.

The Limeil-Brévannes study. Between June 1995 and March 1996, 127 patients with dementia (characterised according to DSM-III-R criteria) were recruited for a clinical study among individuals hospitalised in the long-term care facility of the Department of Gerontology of the Limeil Brévannes Hospital (France).[14] All patients included in this study underwent detailed clinical examination, MMSE score, Hachinski ischaemic score and CT scan. Dementia was identified according to DSM-III-R criteria. For controls, we used 190 brains from routine autopsies carried out at the Hospices Civils de Strasbourg (France). Selection of these controls excluded dementia cases. All cases referred to autopsy for neurological pathologies were also excluded. In order to ensure the exclusion of subclinical AD-affected cases, neuropathological criteria were also applied to exclude tissues from individuals presenting with a Braak stage greater than 2 ($n=59$).

The Northern AD case–control study. The French AD and control samples were Caucasian (AD cases $n=609$, controls $n=567$).[15] Clinical diagnosis of probable AD was established according to the DSM-III-R and NINCDS-ADRDA criteria. Caucasian controls were defined as subjects without DMS-III-R dementia criteria and with integrity of their cognitive functions. Controls were recruited in retirement homes or from electoral rolls (altruistic volunteers). Five AD cases were not successfully genotyped for rs2300403 SNPs.

The UK AD case–control study. All AD cases were Caucasians ($n=349$), ascertained from two UK centres, the central belt of Scotland and Greater Manchester.[16] Diagnoses of definite or probable AD

were established according to DSM-III-R and NINDCS-ADRDA criteria. The proportion of definite AD cases was 30%. The proportion of AD cases with family history was 20%. The Caucasian controls ($n = 398$) were collected from the same geographical areas as the AD patients and were defined as subjects without DSM-III-R dementia criteria and with full integrity of their cognitive functions.

The Pittsburgh study. The Pittsburgh Study: 200 late-onset AD were from the University of Pittsburgh Alzheimer's Disease Research Center (ADRC). Clinical diagnoses of the patients were made according to the NINCDS/ADRDA criteria. The ADRC follows a standard evaluation protocol, which includes medical history, general medical and neurological examinations, a psychiatric interview, neuropsychological testing and a magnetic resonance imaging scan. Age- and sex-matched 200 controls were recruited from the same Western Pennsylvania region as the cases and were determined to be cognitively intact following extensive clinical examination. Three controls and three AD cases were not successfully genotyped for the rs2300403 SNP.

Brain tissues

AD brains were obtained at autopsy from 114 patients with early- and late-onset sporadic AD accessioned from Greater Manchester region of United Kingdom during years 1986 and 2001 (mean age at onset = 65.9 ± 10.3 years; mean age at death = 73.1 ± 9.1 years; 51% men). All patients were Caucasian. Pathological diagnoses were made in accordance with CERAD Neuropathological Criteria for AD. All patients were at Braak stages 5 or 6 at time of death. Tau load in frontal cortex was determined by image analysis in 86/114 patients after immunostaining for phosphorylated tau using a standard procedure with monoclonal antibody AT8 (Innogenetics, Gent, Belgium) as primary antibody.[17]

As controls, we used 118 control brains presenting Braak stages less than 3, selected from an initial set of the 190 brains obtained from routine autopsies carried out at the Hospices Civils de Strasbourg (France, see description in the Limeil-Brevanne study in the Supplementary material).

dHPLC and genotyping

All exons, intron/exon boundaries, entire intron 2 and proximal promoter of the *S100β* gene were screened in 32 healthy individuals from the ELDNOR study (see population description) for sequence variation using denaturing high-performance liquid chromatography (dHPLC) (Supplementary appendix, Table S4). All variants identified by dHPLC were confirmed by sequencing. When already referenced, the name of the SNP was indicated (see Supplementary material, Table S5). The genotyping of 25 SNPs was mainly performed by enzymatic digestion following PCR amplification (see Supplementary material, Table S4). A subpopulation of 254 individuals was used for LD estimation between the S100*β* SNPs (see Supplementary material, Figure S6) and for the systematic screening of their impact on the cognitive performances (MMSE > 24, see population description for more details).

The rs2300403 SNP was genotyped (i) by *Nla*III digestion following PCR amplification using 5′-actctgaaccattcacggtg-3′ and 5′-gtctctcaccaagcccatt-3′ oligonucleotides (INSERM U744) or (ii) by Taqman technology (Genoscreen, see Supplementary material, Table S5). The 254 individuals of the ELDNOR population were genotyped using the two technologies and no discrepancies were observed.

Detection of S100β and S100β2 mRNAs

New splicing events were detected by a computational procedure that analysed the genomic-expressed sequence tag (EST)-mRNA pair-wise alignments. Total RNA from brain or lymphocytes of healthy individuals were used to detect S100*β* and S100*β*2 mRNA as well as purified lymphocytes from fresh blood of Hamlyn's cercopithec. The lymphocytes were cultivated for 72 h in the presence of 0.1% phytohaemagglutinin. Total RNA from lymphocytes was extracted using RNeasy Mini kit (Qiagen, Courtaboeuf, France) and DNase treatment. Reverse transcription (RT) was performed from 30 ng of total RNA treated by DNase before cDNA synthesis. Characterisation of the new exon used four different oligonucleotide sets. Expression of the normal S100*β* mRNA was checked (see Supplementary material, Table S6). Amplification products were directly sequenced using the Taq dye terminator sequencing kit (Applied Biosystems, Foster City, CA, USA). The corresponding mRNA sequences were used to determine the complete mRNA isoform sequences (see Supplementary material, Figure S6). The potential S100*β*2 protein sequence was determined using the Traduc software (http://www.infobiogen.fr/services/analyseq/cgi-bin/traduc_in.pl).

Inter-species comparisons

Genomic DNA was extracted from fresh blood of Pan troglodyte's chimpanzees, Hamlyn's cercopithec (*Cercopithecus hamlynis*), Roloway's cercopithec (*Cercopithecus diana roloway*) and catta lemurian. The *S100β* gene has been sequenced using the human primer sets described in Table S3 (see Supplementary material). The intron 2 sequence was obtained from chimpanzees and cercopithec and compared with the human sequence using the dialign software (http://bioweb.pasteur.fr/seqanal/interfaces/dialign2-simple.html). No amplification was obtained from the lemurian genomic DNA using the oligonucleotides for the amplification of the *S100β2* exonic sequence. Finally, the *S100β2* exonic sequence was searched using the BLAST software (http://www.ncbi.nlm.nih.gov/BLAST/) in the Mus Musculus genome.

Real-time RT-PCR

Total RNA from brains of 84/114 AD patients (74%) and 92/118 controls (78%) was extracted from frozen brain tissue using phenol/chloroform protocol (TRI-zol reagent, Invitrogen, Carlsbad, CA, USA) The quality of total RNA was assessed using Agilent 2100 Bioanalyser and the ratio of ribosomal RNA 28S/18S systematically estimated using the Agilent 2100 Bioanalyser bio-sizing software (range from 0.0 to 1.8 in brains). RT was performed from 30 ng of total RNA and expression was measured by real-time PCR using Taqman technology to co-amplify cDNA from the *S100β* and *β-actin* genes or from the *S100β* 2 and *β-actin* genes, as described by the Supplier (Applied Biosystems, Foster City, CA, USA).

Statistical analyses

The SAS software release 8.02 was used for statistical analyses (SAS Institute, Cary, NC, USA). Pair-wise LD coefficients were estimated in a subset of the ELDNOR population with a MMSE > 24 (254 individuals, see dHPLC and genotyping section) using the GOLD software (http://www.sph.umich.edu/csg/abecasis/GOLD) (see Supplementary material, Table S8).

We used Akaike Information Criterion (AIC) to determine the best-fitting genetic model (dominant, co-dominant or recessive).[18,19] The model with the lowest AIC reflects the best balance of goodness-of-fit and parsimony. We consequently coded the genotypes of the rs2300403 polymorphism as a dummy variable according to the hypothesis of a recessive model, that is, GG versus AG + AA genotype.

The impact of the *S100β* polymorphisms on MMSE was first evaluated using a non-parametric Wilcoxon test in the ELDNOR subpopulation ($n = 254$, see description in the population section). Following the extension of the rs1051169, rs2839355 and rs2300403 SNPs analyses to the whole ELDNOR study, the impact of these SNPs on MMSE was analysed using a general linear model (MMSE analysed as a continuous variable) to estimate the effect of potential confounders such as age, gender or *APOE* status. These complementary analyses did not affect the results (data not shown). As the MMSE scores did not follow a normal distribution, log-transformation was also performed. This complementary analysis did not modify the results (data not shown). Extended haplotype frequencies of the different markers were estimated using the Thesias software, as well as the impact of these haplotypes on MMSE. The objective of the thesias software is to performed haplotype-based association analysis in unrelated individuals. This programme is based on the maximum-likelihood model described in Stram *et al*.[20] and is linked to the SEM algorithm.[21]

The impact of the rs2300403 SNP in the Leiden-85 +, Danish or combined populations was analysed accordingly to the same schema than the one described in the ELDNOR study (see above). The general linear model was adjusted for age, gender, centre and *APOE* status and centre when necessary. Furthermore, as the MMSE scores did not follow a normal distribution and to limit consequent potential biases, the impact of the rs2300403 *S100β* SNP on the cognitive function was also qualitatively analysed by coding the MMSE score as a discontinuous variable using the median MMSE score as a cutoff (two categories: MMSE ⩽ 25 and MMSE > 25). A multiple logistic regression model was then performed adjusted for age, gender, centre and *APOE* status. For all MMSE analyses, individuals exhibiting a MMSE less than 10 were systematically excluded from these analyses.

To study the impact of the rs2300403 SNP on dementia and AD, a univariate analysis was first performed using Pearson's χ^2 test. Before pooled analyses, homogeneity between populations was tested using Breslow–Day computation.[22] The impact of the rs2300403 SNP on dementia or AD was then estimated by multiple logistic regression models, adjusted for age, gender, *APOE* status and centre when necessary. Interactions among rs2300403, *APOE*, gender or age variables were tested in logistic regression models.

RNA is notoriously labile, highly sensitive to RNAse and as a consequence, post-mortem delay can strongly affect the results when studying mRNA expression. However, to use only post-mortem delay as a potential confounder does not allow correcting for history of storage, transportation and/or extraction, leading to further autolyses. To take into account these biases, quality of total RNA and degradation was assessed using ratio of ribosomal RNA 28S/18S measured by the Agilent 2100 bionalyser and bio-sizing software, a technology allowing estimation of RNA degradation (and particularly RNAse action).[23–25] An analysis of covariance (adjusted for age and gender) using a general linear model for comparison of *S100β* or *S100β2* mRNA amounts between AD cases and controls was performed (mRNA level data was log transformed to normalise distributions). The effect of the rs2300403 SNP on *S100β* and *S100β2* mRNA levels was analysed with a non-parametric Wilcoxon test. We also adjusted for mRNA degradation. The results were not modified following this additional analysis (data not shown). Correlation between *S100β*, *S100β2* mRNA levels and/or tau load was tested using non-parametric Spearman statistic. As this level of total RNA degradation (estimated by rRNA 28S/18S ratio using Agilent 2100 Bioanalyser) may be a potential confounder, standardised residuals of linear regression analysis between degradation and *S100β*/*S100β2*/*β-actin* ratio were also held for correlations between tau loads and mRNA amount.[26] These complementary analyses did not affect the results as well as complementary adjustment for age and gender (data not shown).

Results

By bio-informatics, we detected a new splicing event from the Unigene database (http://www.ncbi.nlm.

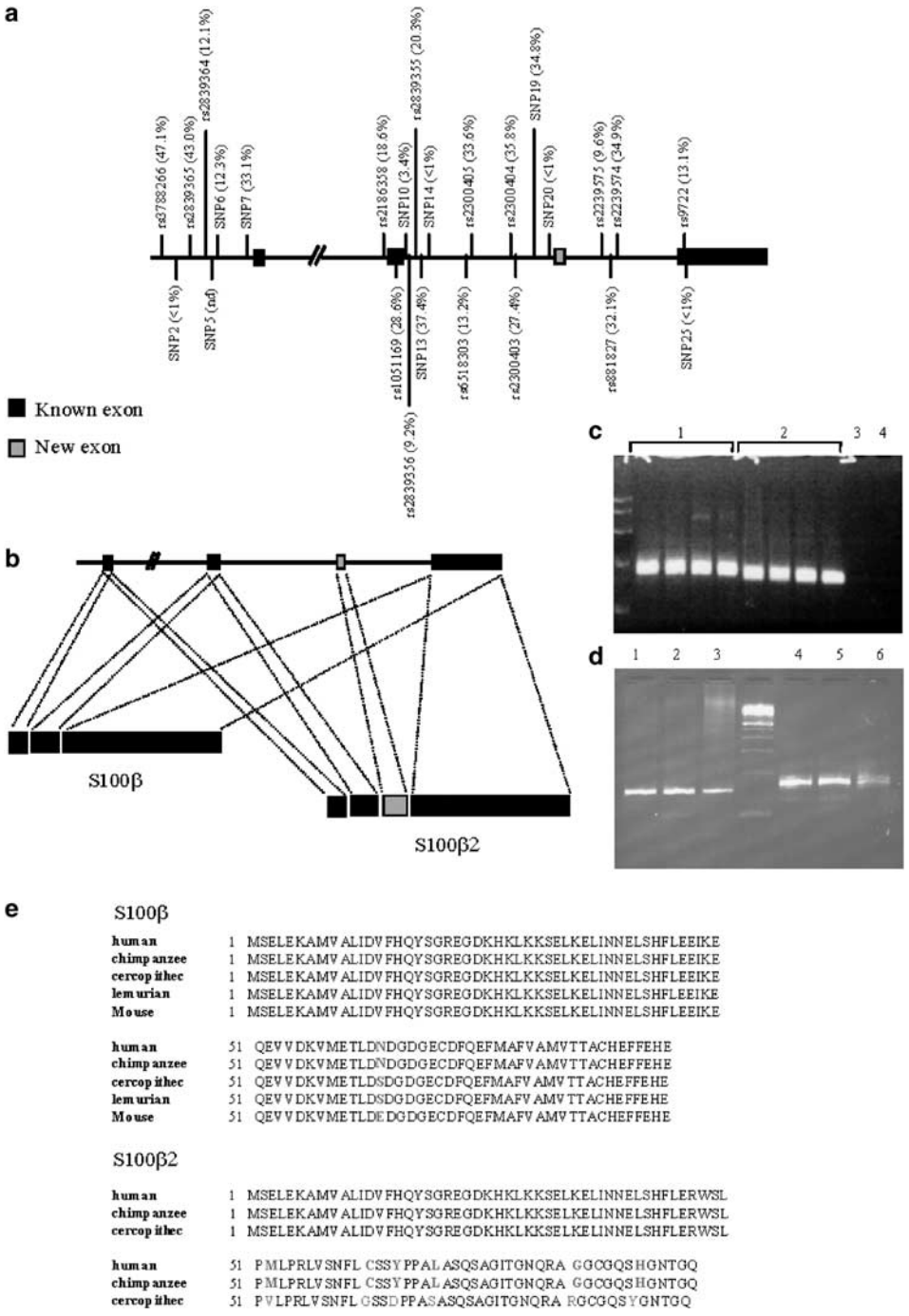
nih.gov/entrez; accession number Hs.422181). This event was observed in nine of the 285 (3.2%) *S100β* ESTs/mRNAs available in this database (Figure 1b). We confirmed the existence of this new exon of 94 bp by RT-PCR from human brain tissues or lymphocytes (Figure 1c,d) and we called this mRNA isoform *S100β2*. The *S100β2*-predicted protein sequence was 95 amino acids (aa) long, differing from S100*β* isoform by the last 48 aa (Figure 1e). The S100*β*2 exonic sequence was found in the genome of chimpanzees and cercopithecs, but not in that of mice and lemurians (Figure 1e). From this observation, the promoter, exons, intron/exon boundaries and the entire intron 2 (containing the new splicing event) were then screened for variants by dHPLC and then sequenced in 32 healthy subjects. Twenty-five SNPs were identified. Nine had not been described previously.

These 25 SNPs (with allele frequencies varying from <1 to 47.1%; Figure 1a) were analysed with respect to cognitive performance in 254 healthy individuals with a MMSE >24, ascertained from the ELDNOR population-based study. Three of these SNPs (rs1051169, rs2839355 and rs2300403) were associated with lower cognitive performances (data not shown). Covariate analyses of these three SNPs over the whole ELDNOR population showed rs2300403 SNP to be the one most strongly associated with low cognitive performance (see Supplementary material, Table S1). Further analyses confirmed this SNP association in a Danish twin population (Table 1). A similar trend was observed in the Leiden 85-plus population-based study (Table 1). The association of the rs2300403 GG genotype was confirmed when the three populations were pooled (Table 1) and this association seemed to be independent of a dementia or a predementia status (MMS >24, ΔMMSE $=-0.5$, $P=0.002$; MMSE >27, ΔMMSE $=-0.3$, $P=0.01$). The MMSE score was finally dichotomised using the median score as a cutoff ($\leqslant 25$ and >25 in the combined population, see Materials and methods). Individuals bearing the GG genotype had a 70% increased risk of low cognitive performance compared with the AA carriers (OR = 1.7, 95 CI% (1.2–2.1), $P=0.0005$; Table 2). This observation was consistent whatever the studied population (ORs ranging from 1.5 to 2.8, data not shown). All these data indicated that the rs2300403 SNP is associated with lower cognitive performance in the elderly (age at examination in the pooled population: 83.3 ± 7.3 years). No significant statistical interaction between gender, age, *APOE* status and the risk of low cognitive performance was detected (data not shown).

Next, we investigated the potential impact of the rs2300403 SNP on the risk of dementia in three independent populations (Leiden-85+, ELDNOR and Limiel-Brevannes). The rs2300403 GG genotype was consistently associated with an increased risk of developing dementia in the three studies (ORs ranging from 1.8 to 3.0; OR = 2.1, $P=0.001$ in the combined population; Table 3). We systematically searched for potential statistical interaction between gender, age, APOE status and the risk of developing dementia. In these studies, the association of the rs2300403 GG genotype was dependent on gender and was only associated with an increased risk in women (ORs ranging from 2.2 to 5.7; OR = 2.5, $P=0.0006$ in the combined population; gender interaction $P=0.02$; Table 3 and Supplementary material, Figure S1). All these data suggested that the rs2300403 SNP was associated with the risk of dementia in the elderly (age at examination in the pooled population: 84.7 ± 6.8 years), especially in women.

Lastly, we assessed the impact of the rs2300403 polymorphism in AD, the most common cause of dementia, in three different case–control studies. In American and UK populations, the rs2300403 GG genotype was associated with an increased risk of AD (OR = 2.0, $P=0.05$ and OR = 1.7, $P=0.03$, respectively, Table 3). In a northern French study, a similar trend was observed (OR = 1.4, $P=0.10$, Table 3). Analysis of the pooled populations strongly supported a deleterious effect of the GG genotype (OR = 1.7, $P=0.0009$, Table 3). We again searched for potential statistical interaction between gender, age, APOE status and the risk of developing AD. In these studies, the association of the rs2300403 GG genotype was again dependent on gender. As observed for dementia, the rs2300403 GG genotype was consistently associated with an increased risk of AD only in women (ORs ranging from 1.6 to 2.2; OR = 1.9, $P=0.003$ in the combined population; gender interaction $P=0.05$; Table 3 and Supplementary material, Figure S2). An interaction between age and the association of the rs2300403 GG genotype with AD was jointly detected ($P=0.008$). Supporting this interaction, we observed that patients bearing the GG genotype had a higher age at onset compared with the AA carriers (69.9 ± 9.1 versus 68.3 ± 9.3, $P=0.04$), this effect being more pronounced in women (71.5 ± 8.7 versus 69.19.7, $P=0.02$; gender interaction $P=0.02$). Consistently, in older AD patients (using the median age at examination of the different populations as a cutoff), the rs2300403 GG genotype was associated with an increased risk of developing AD (OR = 2.8, $P<0.0001$, Table 3), whereas no effect was detected in the younger patients (OR = 1.1, not significant, Table 3). The rs2300403 SNP may therefore increase the risk of developing AD in elderly persons (age at examination in the pooled population: 81.1 ± 5.0 years). This observation was again consistent whatever the population (ORs ranging from 2.4 to 3.1; see Supplementary material, Figure S3) and with our observation in dementia in the elderly (age at examination in the pooled population in the dementia study: 84.7 ± 6.8 years). As expected, the highest risk of developing AD associated with the rs2300403 GG genotype was observed in the oldest women (OR = 3.6, 95 CI% (1.9–6.8), $P<0.0001$).

As the rs2300403 SNP is located in the intron upstream from the new 100*β*2 isoform, we postulated



e

S100β

human	1	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKE
chimpanzee	1	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKE
cercopithec	1	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKE
lemurian	1	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKE
Mouse	1	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLEEIKE
human	51	QEVVDKVMETLDNDGDGECDFQEFMAFVAMVTTACHEFFEHE
chimpanzee	51	QEVVDKVMETLDNDGDGECDFQEFMAFVAMVTTACHEFFEHE
cercopithec	51	QEVVDKVMETLDSDGDGECDFQEFMAFVAMVTTACHEFFEHE
lemurian	51	QEVVDKVMETLDSDGDGECDFQEFMAFVAMVTTACHEFFEHE
Mouse	51	QEVVDKVMETLDEDGDGECDFQEFMAFVAMVTTACHEFFEHE

S100β2

human	1	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLERWSL
chimpanzee	1	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLERWSL
cercopithec	1	MSELEKAMVALIDVFHQYSGREGDKHKLKKSELKELINNELSHFLERWSL
human	51	PMLPRLVSNFLCSSYPPALASQSAGITGNQRAGGCGQSHGNTGQ
chimpanzee	51	PMLPRLVSNFLCSSYPPALASQSAGITGNQRAGGCGQSHGNTGQ
cercopithec	51	PVLPRLVSNFLGSSDPPASASQSAGITGNQRARGCGQSYGNTGQ

Figure 1 Description of the *S100β* gene. (**a**) Position of SNPs within the *S100β* gene. (**b**) Organisation of the new splicing event of the *S100β* gene. (**c**) Amplification of the S100*β*2 isoform using (1) the primer set 3 or (2) the primer set 1 in four AD brain samples. (3) and (4) were respectively negative controls for primer set 3 and primer set 1 amplifications. (**d**) Amplification of the S100*β* isoform (1–3) and S100*β*2 isoform (4–6; primer set 2) by RT-PCR from human brain (1 and 4), human (2 and 5) and Cercopithec (3 and 6) lymphocytes. PCR products were systematically sequenced to confirm the specific amplification of the mRNAs of interest. (**e**) Comparison of the putative protein sequences based on the different S100*β* exon sequences. The mouse protein sequence was obtained from the NCBI website (BC061178). The additional exon was successfully amplified only from human, chimpanzees (two different samples) and cercopithecs (two species: Hamlyn's and Roloway's cercopithecs) genomic DNA. Variations in aa sequences are indicated in purple. The sequence of S100*β* protein differed between the different species studied by only one residue at codon 63, whereas the putative S100*β*2 protein may differ by six aa between cercopithecs and humans (chimpanzee possessed the same S100*β*2 aa sequence as humans). These data indicated that the retention of this exon may be recent in terms of evolution.

Table 1 Impact of the rs2300403 SNP on cognitive performance in three populations

	ELDNOR		*Denmark*		*Leiden*		*Combined*	
	n	*MMS*	n	*MMS*	n	*MMS*	n	*MMS*
(a)								
AA	399	21.2±5.3	233	25.7±3.7	462	25.4±4.6	1094	23.9±5.1
AG	331	21.2±5.1	229	26.0±3.1	488	25.0±4.9	1048	24.0±5.0
GG	85	19.6±4.9	51	24.9±3.6	99	24.7±4.6	235	22.9±5.1
P		0.03		0.11		0.24		0.002
(b)								
AA+AG	730	21.2±5.2	462	25.8±3.4	950	25.2±4.7	2142	24.0±5.1
GG	85	19.6±4.9	51	24.9±3.6	99	24.7±4.6	235	22.9±5.1
ÄMMS		−1.6		−0.9		−0.5		−1.1
P		0.008		0.04		0.14		0.0005

Table 2 Estimation of the risk of low cognitive performance according to rs2300403 genotype in the combined population

Combined[a]	n	*AA*	*AG*[b]	*GG*[c]
MMSE > 25	1162	554 (0.48)	518 (0.44)	90 (0.08)
MMSE ⩽ 25	1215	540 (0.44)	530 (0.44)	145 (0.12)

Abbreviation: MMSE, mini mental-state examination.
[a]$P = 0.002$.
[b]OR (AG versus AA) = 1.0, 95 CI% (0.9–1.2), $P = 0.60$.
[c]OR (GG versus AA) = 1.7, 95 CI% (1.2–2.3), $P = 0.0005$.

that its deleterious association with cognitive performance and dementia could be related to the expression of this isoform. We first measured the levels of *S100β* and *S100β2* mRNAs in the brain of 84 AD cases and 90 controls by semiquantitative RT-PCR. We observed a strong inverse relationship between the level of *S100β* mRNA and the level of *S100β2* mRNA either in the brain of AD cases and controls ($P < 0.0001$; see Supplementary material, Figure S4). This observation suggested a potential regulatory mechanism controlling the balance of synthesis between the two isoforms. A 37% decrease in *S100β* mRNA ($P < 0.0001$) and a 41% increase in *S100β*2 mRNA ($P = 0.0007$) were measured in AD compared with controls (Figure 2a). These decreases in *S100β* expression, and increases in S100β2 expression, were greater in women (Figure 2a), in accordance with previous epidemiological data. As postulated, we observed that the GG genotype was associated with a 79% increase in S100β2 mRNA levels ($P = 0.01$) compared with cases bearing AA and AG genotypes in AD brains ($n = 84$, Figure 2b). Consistent with previous results, this increase was higher in women (Figure 2b). In control brains, the rs2300403 GG genotype was not associated with an increase in *S100β2* mRNA level even if a trend may be observed (+20%, data not shown). Therefore, the rs2300403 polymorphism may modulate mechanisms controlling *S100β2* mRNA synthesis, especially in the brain of women suffering from AD.

Discussion

We analysed 25 SNPs in the *S100β* locus and our findings suggest that at least the rs2300403 SNPs in S100β increases the risk of low cognitive performance, dementia and AD, possibly by favouring a splicing event leading to the S100β2 isoform in the brain.

All the *P* values we reported for the genetic analyses are not corrected for multiple comparisons. However, we observed a consistent and highly significant association of the *S100β* rs2300403 SNP with low cognitive performances using the same recessive model in three independent populations and with the risk of developing dementia, especially in women, in six independent populations. We systematically observed throughout our various experiments, significant and consistent effects. Reinforcing the coherence of our genetic epidemiology data, we observed a significant association of the rs2300403 SNP with a higher level of *S100β2* expression in the brain of AD cases, again especially in women.

However, nonetheless and notwithstanding the significant and consistent effects we have observed, additional genetic studies involving prospective cohorts as well as both family-based and large case–control samples will be required to characterise fully the role of the rs2300403 SNP. Furthermore, we cannot exclude the possibility of LD with causal or more additional pathogenic variants. To evaluate this possibility, we first screened for all SNPs in the entire exon 2, intron 2 and exon 3 of the *S100β* gene, susceptible to modulate the retention of the new *S100β2* exon. We finally restricted our analyses to three SNPs (rs1051169, rs2839355 and rs2300403), exhibiting a significant association with low cognitive performances in a first screening. In addition with covariate analyses (see Supplementary material, Table

Table 3 Association of the rs2300403 SNP with dementia or AD by study according to gender or age (%)

Study		*rs2300403,* n *(%)*					*OR (95% CI)*	
		n	*AA*	*AG*	*GG*	P	*AG versus AA*	*GG versus AA*
Dementia								
ELDNOR	Cases	203	98 (0.48)	84 (0.41)	21 (0.10)	0.05	1.0 (0.6–1.6)	2.9 (1.0–8.3)
	Controls	146	80 (0.55)	61 (0.42)	5 (0.03)		NS	0.04
Limeil-Brevannes	Cases	127	62 (0.49)	44 (0.35)	21 (0.17)	0.03	1.1 (0.6–1.8)	3.0 (1.3–7.4)
	Controls	131	72 (0.49)	51 (0.41)	8 (0.06)		NS	0.02
Leiden 85-plus	Cases	206	75 (0.36)	109 (0.53)	22 (0.11)	0.06	1.5 (1.0–2.1)	1.8 (1.0–3.3)
	Controls	438	202 (0.46)	202 (0.46)	34 (0.08)		0.04	0.06
Combined	Cases	536	235 (0.44)	237 (0.44)	64 (0.12)	0.002	1.2(0.9–1.5)	2.1 (1.3–3.3)
	Controls	715	354 (0.49)	314 (0.44)	47 (0.07)		NS	0.001
Combined men	Cases	122	63 (0.52)	50 (0.41)	9 (0.07)	NS	0.9 (0.5–1.3)	1.2 (0.5–3.1)
	Controls	266	124 (0.47)	126 (0.47)	16 (0.06)		NS	NS
Combined women	Cases	414	172 (0.42)	187 (0.45)	55 (0.13)	0.001	1.2(1.0–1.8)	2.5 (1.5–4.2)
	Controls	449	230 (0.51)	188 (0.42)	31 (0.07)		0.06	0.0006
AD								
France	Cases	604	282 (0.47)	260 (0.43)	62 (0.10)	NS	0.9 (0.6–1.5)	1.4 (0.9–2.2)
	Controls	567	279 (0.49)	245 (0.43)	43 (0.08)		NS	0.10
UK	Cases	349	167 (0.48)	133 (0.38)	49 (0.14)	0.04	1.0 (0.7–1.4)	1.7 (1.1–2.8)
	Controls	398	205 (0.52)	160 (0.40)	33 (0.08)		NS	0.03
USA	Cases	197	78 (0.40)	90 (0.46)	29 (0.15)	NS	1.2 (0.8–1.9)	2.0 (1.0–4.0)
	Controls	197	89 (0.45)	86 (0.44)	22 (0.11)		NS	0.05
Combined	Cases	1150	527 (0.46)	483 (0.42)	140 (0.12)	0.009	1.0 (0.9–1.3)	1.7 (1.2–2.3)
	Controls	1162	573 (0.49)	491 (0.43)	98 (0.08)		NS	0.0009
Combined men	Cases	452	205 (0.45)	202 (0.45)	45 (0.10)	NS	1.2 (0.9–1.6)	1.3 (0.8–2.0)
	Controls	526	266 (0.50)	214 (0.41)	46 (0.09)		NS	NS
Combined women	Cases	689	318 (0.46)	277 (0.40)	94 (0.14)	0.006	1.0 (0.8–2.0)	1.9 (1.3–2.8)
	Controls	636	307 (0.48)	277 (0.44)	52 (0.08)		NS	0.003
Combined age $\leqslant 72$	Cases	571	238 (0.48)	210 (0.42)	49 (0.10)	NS	1.1 (0.9–1.4)	1.1 (0.7–1.6)
	Controls	642	296 (0.50)	239 (0.41)	52 (0.09)		NS	NS
Combined age > 72	Cases	526	211 (0.46)	183 (0.40)	62 (0.14)	0.0004	1.0 (0.7–1.3)	2.8 (1.7–4.6)
	Controls	505	188 (0.50)	166 (0.44)	24 (0.06)		NS	< 0.0001

Abbreviations: AD, Alzheimer's disease; NS, not significant; SNP, single nucleotide polymorphism.
ORs were adjusted on age, gender, *APOE* status and centre when necessary.

S1), we also performed haplotype analysis using the rs1051169, rs2839355 and rs2300403 SNPs in the whole ELDNOR study. We investigated the association of the resulting haplotypes with low cognitive performances using the Thesias software.[21] This haplotype analysis assumes a genetic additive model, whereas the most pertinent model in our case is a recessive one.[21] Nevertheless, despite this limitation, haplotypes bearing the rs2300403 G allele are systematically associated with a lower cognitive performance (see Supplementary material, Table S7), supporting the initial results we observed in the covariate analysis (see Supplementary material, Table S1). Altogether, these data support a main effect of the rs2300403 SNP. However, only functional studies to evaluate the impact of each SNP on the S100β2 synthesis may help to characterise fully the causal SNP(s). For instance, the systematic gender effect we observed, may indicate a hormone-mediated RNA splicing. Such mechanisms have been already described with the oestrogen receptors.[27–29]

It is not clear how a modification of expression of the *S100β* and *S100β2* mRNA ratio may favour biological processes leading to low cognitive performance and dementia. A number of observations point to a beneficial physiological role for S100β. This suggests that an increase in expression of S100β in response to injury might be associated with injured neurons repair, cellular debris clean-up and further damage resistance (for review, see Donato[30]). On the other hand, it has been suggested that chronically increased S100β levels in the adult brain can be deleterious. For instance, the elevated tissue level of S100β in AD brain correlate with the density of neuritic plaques[31,32] and with the density of dystrophic neuritis within plaques.[33] Peskind *et al.*[34] reported that the cerebro-spinal fluid (CSF) S100β concentration was significantly higher in 46 mild/moderate stage AD subjects compared with 25 healthy controls. This observation may appear to be in contradiction with our observation of a 47% decrease in *S100β* mRNA level in the brain of AD cases.

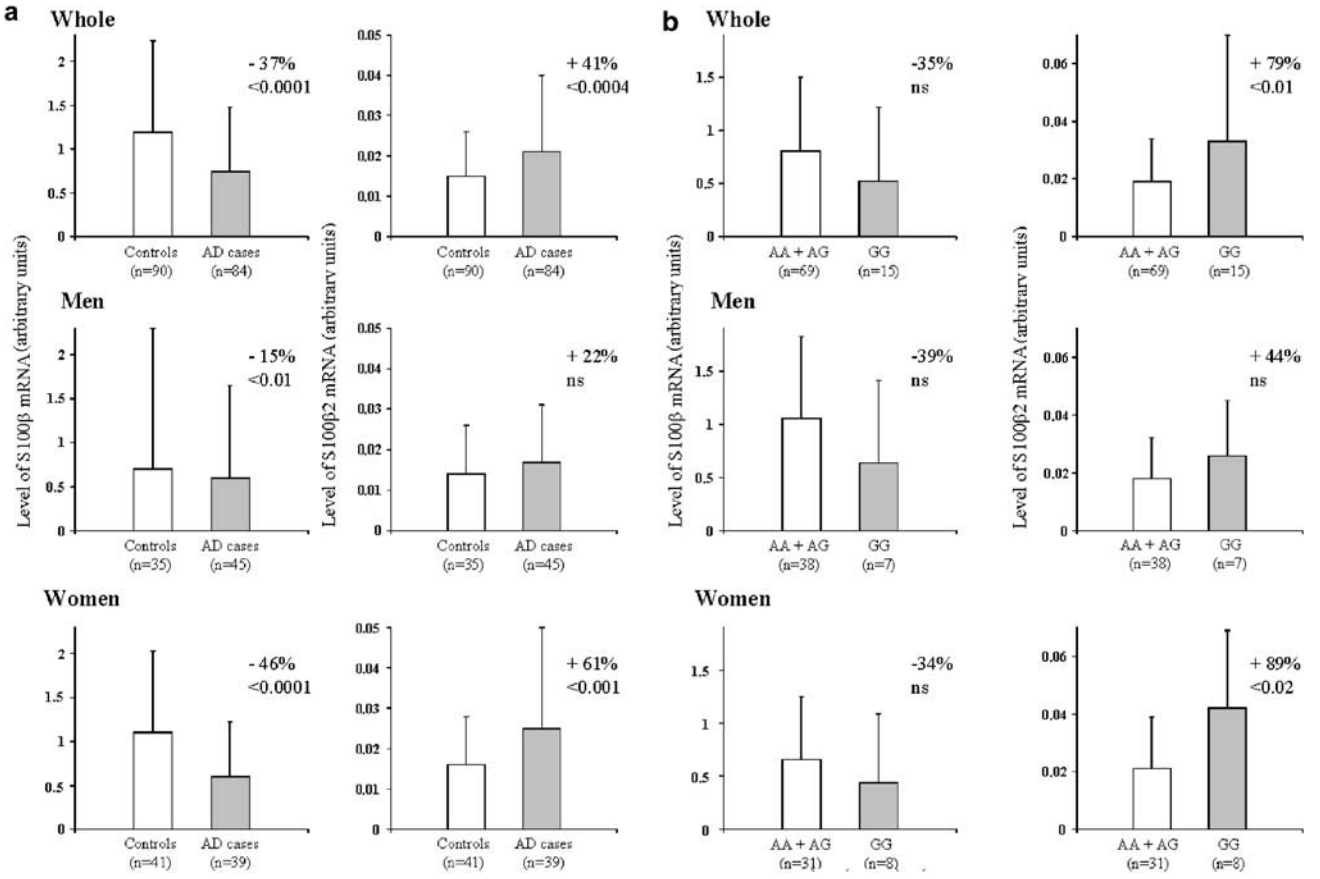


Figure 2 (**a**) Comparison of S100β and S100β2 expression levels (means $\pm$ s.e.m.) in AD and control brains. (**b**) Association between the rs2300403 genotypes and the mRNA expression level of S100β and S100β2 (means $\pm$ s.e.m.). Non-parametric Wilcoxon tests were used.

However, in the same paper, Peskind *et al.*[34] also noticed a nonsignificant 40% decrease in CSF S100β concentration in 22 more advanced-stage AD subjects compared with controls. Interestingly, they also observed a significant positive correlation in all AD subjects between CSF S100β and MMSE scores and a significant negative correlation between CSF S100β and CDR stage.[34] At this stage, it is interesting to note that we observed that in AD brains, the total amount of hyperphosphorylated tau proteins was weakly increased when (i) the expression level of S100β decreased ($P = 0.03$) or (ii) the expression level of S100β2 increased ($P = 0.009$, Figure S7). As AD is a multifactorial and heterogeneous disease, it is plausible that numerous genes may modulate tau loads in the brain of AD cases. As a consequence, the observation that 10% of variation in tau protein levels may be attributable to S100β2 gene expression, is relevant. Even if our results are only observational and do not allow to establish a causative link, the correlations of the S100β or S100β2 expression levels with tau level are of interest as neurofibrillary degeneration may constitute a pathological substrate for memory loss in normal ageing, mild cognitive impairment and AD.[35,36] Interestingly, in addition to activation of cytokines, it has been described that S100β may also play a direct role in tau-related pathways that are associated with neurodegeneration: S100β may be a modulator of tau phosphorylation and any changes in their interaction could be a factor in the AD-related hyperphosphorylation.[37–39]

At this stage, to determine the biological activities associated with the *S100β2* mRNA isoform may clearly help to understand better the role of S100β in the dementia process. Indeed, it is unlikely that this mRNA isoform results from an aberrant splicing event as this one appears to be relatively frequent, as observed in EST database (3.2%) and brains of AD cases (5.8% in average of the total mRNAs produced from the *S100β* locus and up to 50% in some cases). Two possibilities may be hypothesised: (i) the mRNA isoform may be translated into a protein which differs by its last 48 aa from S100β. Research for protein homologies by BLAST or specific domain in the new C-terminal domain were unsuccessful and to estimate a possible gain of function of S100β2 compared with S100β was, as a consequence, not possible. However, a potential loss of function may be evaluated. The neurotrophic activities of S100β are dependent on a disulphide-linked dimeric form of the protein.[40–42]

The number of cysteine residues are not modified between the S100β and S100β2 isoforms (at codon 68 and 84 and at codon 62 and 84, respectively), suggesting a conserved ability to form disulphide-linked dimer. However, the S100β C-terminal domain is important for certain S100β/target protein interactions.[43–45] With this background, one possibility would be that S100β2 interacts with S100β, modulating the neurotrophic biological activities normally associated with the S100β homodimer; (ii) the S100β2 mRNA isoform is finally not translated and is produced to modulate the *S100β* mRNA pool, available for S100β protein production. Such hypothesis could be supported by the strong correlation existing between the S100β and S100β2 mRNA levels in the brain of AD cases. Thus, the next step of this work will be developing the tools necessary to answer the biological relevance of the *S100β2* mRNA isoform. More extensive functional testing of S100β2 will be required to fully characterise the role of the S100β2 mRNA isoform in the physiopathological processes leading to AD and dementia, and as a consequence to better determine the role of S100β in these processes.

In conclusion, our results suggest that genetic variations in *S100β* are associated with increased risk of lower cognitive performance, dementia and AD. These effects are more pronounced in women and increase with age. This interaction with age reinforces the hypothesis that S100β plays a role in the continuum from cognitive impairment to dementia. Furthermore, this effect may be mediated by a splicing event leading to a new primate-specific *S100β* mRNA isoform. Furthermore, as the S100β protein has already been implicated in traumatic, ischaemic and inflammatory brain damage as well as in neurodegenerative and psychiatric disorders, our discovery may have major implications for many human brain pathologies.[46–48]

Acknowledgments

We thank Dr Walter Miller for providing DNA samples from chimpanzee. This work was supported by INSERM, the Pasteur Institute of Lille, Genoscreen and by research grants from the National Institute on Aging (Grant NIA-PO1-AG08761).

References

1 Drachman DA. If we live long enough, will we all be demented? *Neurology* 1994; **44**: 1563–1565.
2 Hardy J. Amyloid, the presenilins and Alzheimer's disease. *Trends Neurosci* 1997; **20**: 154–159.
3 Farrer LA, Cupples LA, Haines JL, Hyman B, Kukull WA, Mayeux R *et al.* Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis. APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA* 1997; **278**: 1349–1356.
4 Brayne C, Harrington CR, Wischik CM, Huppert FA, Chi LY, Xuereb JH *et al.* Apolipoprotein E genotype in the prediction of cognitive decline and dementia in a prospectively studied elderly population. *Dementia* 1996; **7**: 169–174.
5 Kehoe P, Wavrant-De Vrieze F, Crook R, Wu WS, Holmans P, Fenton I *et al.* A full genome scan for late onset Alzheimer's disease. *Hum Mol Genet* 1999; **8**: 237–2345.
6 Van Eldik LJ, Wainwright MS. The Janus face of glial-derived S100B: beneficial and detrimental functions in the brain. *Restor Neurol Neurosci* 2003; **21**: 97–108.
7 Baudier J, Cole RD. Interactions between the microtubule-associated tau proteins and S100β regulate tau phosphorylation by the Ca2 + /calmodulin-dependent protein kinase II. *J Biol Chem* 1988; **263**: 5876–5883.
8 Mrak RE, Griffin WS. Interleukin-1, neuroinflammation, and Alzheimer's disease. *Neurobiol Aging* 2001; **22**: 903–908.
9 Mrak RE, Griffin WS. The role of activated astrocytes and of the neurotrophic cytokine S100B in the pathogenesis of Alzheimer's disease. *Neurobiol Aging* 2001; **22**: 915–922.
10 Connolly Jr ES, Winfree CJ, Rampersad A, Sharma R, Mack WJ, Mocco J *et al.* Serum S100B protein levels are correlated with subclinical neurocognitive declines after carotid endarterectomy. *Neurosurgery* 2001; **49**: 1076–1082.
11 Amouyel P, Richard F, Cottel C, Amant C, Codron V, Helbecque N. The deletion allele of the angiotensin I converting enzyme gene as a genetic susceptibility factor for cognitive impairment. *Neurosci Lett* 1996; **217**: 203–205.
12 Frederiksen H, Gaist D, Bathum L, Andersen K, McGue M, Vaupel JW *et al.* Angiotensin I-converting enzyme (ACE) gene polymorphism in relation to physical performance, cognition and survival – a follow-up study of elderly Danish twins. *Ann Epidemiol* 2003; **13**: 57–65.
13 Van Exel E, Gussekloo J, Houx P, de Craen AJ, Macfarlane PW, Bootsma-van der Wiel A *et al.* Association between high-density lipoprotein and cognitive impairment in the oldest old. *Ann Neurol* 2002; **51**: 716–721.
14 Richard F, Fromentin-David I, Ricolfi F, Ducimetiere P, Di Menza C, Amouyel P *et al.* The angiotensin I-converting enzyme gene as a susceptibility factor for dementia. *Neurology* 2001; **56**: 1593–1595.
15 Rudrasingham V, Wavrant-De Vrieze F, Lambert J-C, Chakraverty S, Kehoe P, Crook R *et al.* Alpha-2 macroglobulin gene and Alzheimer disease. *Nat Genet* 1999; **22**: 17–19.
16 Lambert J-C, Araria-Goumidi L, Myllykangas L, Ellis C, Wang JC, Bullido MJ *et al.* Contribution of APOE promoter polymorphisms to Alzheimer's disease risk. *Neurology* 2002; **59**: 59–66.
17 Hayes A, Thaker U, Iwatsubo T, Pickering-Brown SM, Mann DM. Pathological relationships between microglial cell activity and tau and amyloid beta protein in patients with Alzheimer's disease. *Neurosci Lett* 2002; **331**: 171–174.
18 Akaike H. A Bayesian analysis of the minimum AIC procedure. *Ann Inst Statist Math* 1978; **30**: 9–14.
19 Bozdogan H. Model-selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. *Psychometrika* 1987; **52**: 345–370.
20 Stram DO, Leigh Pearce C, Bretsky P, Freedman M, Hirschhorn JN, Alshuler D *et al.* Choosing haplotype-tagging SNPs based on unphased genotype data using a preliminary sample of unrelated subjects with an example from the multiethnic cohort study. *Hum Hered* 2003; **55**: 27–36.
21 Tregouet DA, Tiret L. Cox proportional Hazards survival regression in haplotype-based association analysis using the stochastic-EM algorithm. *Eur J Hum Genet* 2004; **12**: 971–974.
22 Breslow NE, Day NE, Halvorsen KT, Prentice RL, Sabai C. Estimation of multiple relative risk functions in matched case-control studies. *Am J Epidemiol* 1978; **108**: 299–307.
23 Kuschel M, Muller O, Buhlmann C. Lab-on-a-chip technology. *Innov Pharm Technol* 2002; **9**: 38–45.
24 Ricicova M, Palkova Z. Comparative analyses of Saccharomyces cerevisiae RNAs using Agilent RNA 6000 Nano Assay and agarose gel electrophoresis. *FEMS Yeast Res* 2003; **4**: 119–122.
25 Gomes LI, Silva RL, Stolf BS, Cristo EB, Hirata R, Soares FA *et al.* Comparative analysis of amplified and non-amplified RNA for hybridization in cDNA microarray. *Anal Biochem* 2003; **321**: 244–251.
26 Lambert J-C, Mann D, Harris J, Araria-Goumidi L, Chartier-Harlin MC, Cottel D *et al.* Is there a relation between APOE expression and brain amyloid load in Alzheimer's disease? *J Neurol Neurosurg Psych* 2005; **76**: 928–933.

27 Masuhiro Y, Mezaki Y, Sakari M, Takeyama K, Yoshida T, Inoue K *et al.* Splicing potentiation by growth factor signals via estrogen receptor phosphorylation. *Proc Natl Acad Sci USA* 2005; **102**: 8126–8131.
28 Zhu N, Eghbali M, Helguera G, Song M, Stefani E, Toro L. Alternative splicing of Slo channel gene programmed by estrogen, progesterone and pregnancy. *FEBS Lett* 2005; **579**: 4856–4860.
29 Auboeuf D, Honig A, Berget SM, O'Malley BW. Coordinate regulation of transcription and splicing by steroid receptor coregulators. *Science* 2002; **298**: 416–419.
30 Donato R. Functional roles of S100 proteins, calcium-binding proteins of the EF-hand type. *Biochim Biophys Acta* 1999; **1450**: 191–231.
31 Sheng JG, Mrak RE, Griffin WS. S100beta protein expression in Alzheimer disease: potential role in the pathogenesis of neuritic plaques. *J Neurosci Res* 1994; **39**: 398–404.
32 Sheng JG, Mrak RE, Griffin WS. Glial-neuronal interactions in Alzheimer disease: progressive association of IL-1alpha+ microglia and S100beta+astrocytes with neurofibrillary tangle stages. *Neuropathol Exp Neurol* 1997; **56**: 285–290.
33 Mrak RE, Sheng JG, Griffin WS. Correlation of astrocytic S100 beta expression with dystrophic neurites in amyloid plaques of Alzheimer's disease. *J Neuropathol Exp Neurol* 1996; **55**: 273–279.
34 Peskind ER, Griffin WS, Akama KT, Raskind MA, Van Eldik LJ. Cerebrospinal fluid S100B is elevated in the earlier stages of Alzheimer's disease. *Neurochem Int* 2001; **39**: 409–413.
35 Giannakopoulos P, Herrmann FR, Bussiere T, Bouras C, Kovari E, Perl DP *et al.* Tangle and neuron numbers, but not amyloid load, predict cognitive status in Alzheimer's disease. *Neurology* 2003; **60**: 1495–1500.
36 Guillozet AL, Weintraub S, Mash DC, Mesulam MM. Neurofibrillary tangles, amyloid, and memory in aging and mild cognitive impairment. *Arch Neurol* 2003; **60**: 729–736.
37 Baudier J, Cole RD. Reinvestigation of the sulfhydryl reactivity in bovine brain S100b (beta beta) protein and the microtubule-associated tau proteins. Ca2+ stimulates disulfide cross-linking between the S100b beta-subunit and the microtubule-associated tau(2) protein. *Biochemistry* 1988; **19**: 2728–2736.
38 Sorci G, Agneletti AL, Donato R. Effects of S100A1 and S100B on microtubule stability. An *in vitro* study using triton-cytoskeletons from astrocyte and myoblast cell lines. *Neuroscience* 2000; **99**: 773–783.
39 Yu WH, Fraser PE. S100beta interaction with tau is promoted by zinc and inhibited by hyperphosphorylation in Alzheimer's disease. *J Neurosci* 2001; **2**: 2240–2246.
40 Scotto C, Mely Y, Ohshima H, Garin J, Cochet C, Chambaz E *et al* Cysteine oxidation in the mitogenic S100B protein leads to changes in phosphorylation by catalytic CKII-alpha subunit. *J Biol Chem* 1998; **273**: 3901–3908.
41 Matsui Lee IS, Suzuki M, Hayashi N, Hu J, Van Eldik LJ, Titani K *et al* Copper-dependent formation of disulfide-linked dimer of S100B protein. *Arch Biochem Biophys* 2000; **374**: 137–141.
42 Koppal T, Lam AG, Guo L, Van Eldik LJ. S100B proteins that lack one or both cysteine residues can induce inflammatory responses in astrocytes and microglia. *Neurochem Int* 2001; **39**: 401–407.
43 Lin J, Blake M, Tang C, Zimmer D, Rustandi RR, Weber DJ *et al* Inhibition of p53 transcriptional activity by the S100B calcium-binding protein. *J Biol Chem* 2001; **276**: 35037–35041.
44 McClintock KA, Van Eldik LJ, Shaw GS. The C-terminus and linker region of S100B exert dual control on protein-protein interactions with TRTK-12. *Biochemistry* 2002; **41**: 5421–5428.
45 Frizzo JK, Tramontina AC, Tramontina F, Gottfried C, Leal RB, Donato R *et al* Involvement of the S100B in cAMP-induced cytoskeleton remodeling in astrocytes: a study using TRTK-12 in digitonin-permeabilized cells. *Cell Mol Neurobiol* 2004; **24**: 833–840.
46 Rothermundt M, Ponath G, Arolt V. S100B in schizophrenic psychosis. *Int Rev Neurobiol* 2004; **59**: 445–470.
47 Mrak RE, Griffin WS. Trisomy 21 and the brain. *J Neuropathol Exp Neurol* 2004; **63**: 679–685.
48 Liu J, Shi Y, Tang J, Guo T, Li X, Yang Y *et al.* SNPs and haplotypes in the S100B gene reveal association with schizophrenia. *Biochem Biophys Res Commun* 2005; **328**: 335–341.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)